

JA



ENTERED

PCT

## RAW SEQUENCE LISTING

DATE: 04/08/2002

PATENT APPLICATION: US/09/913,414

TIME: 14:15:05

Input Set : A:\08106-005001.TXT

Output Set: N:\CRF3\04082002\I913414.raw

4 <110> APPLICANT: Smit, John  
 6 <120> TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS POLYPEPTIDES  
 7 FROM FRESHWATER CAULOBACTER  
 10 <130> FILE REFERENCE: 08106-005001  
 12 <140> CURRENT APPLICATION NUMBER: 09/913,414  
 13 <141> CURRENT FILING DATE: 2001-08-13  
 15 <150> PRIOR APPLICATION NUMBER: PCT/CA00/00173  
 16 <151> PRIOR FILING DATE: 2000-02-21  
 18 <150> PRIOR APPLICATION NUMBER: CA 2261186  
 19 <151> PRIOR FILING DATE: 1999-02-19  
 21 <160> NUMBER OF SEQ ID NOS: 7  
 23 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 82  
 27 <212> TYPE: PRT  
 28 <213> ORGANISM: Caulobacter crescentus  
 30 <400> SEQUENCE: 1  
 31 Ala Phe Gly Ala Ala Val Thr Leu Gly Ala Ala Ala Thr Leu Ala Gln  
 32 1 5 10 15  
 33 Tyr Leu Asp Ala Ala Ala Gly Asp Gly Ser Gly Thr Ser Val Ala  
 34 20 25 30  
 35 Lys Trp Phe Gln Phe Gly Gly Asp Thr Tyr Val Val Val Asp Ser Ser  
 36 35 40 45  
 37 Ala Gly Ala Thr Phe Val Ser Gly Ala Asp Ala Val Ile Lys Leu Thr  
 38 50 55 60  
 39 Gly Leu Val Thr Leu Thr Thr Ser Ala Phe Ala Thr Glu Val Leu Thr  
 40 65 70 75 80  
 41 Leu Ala  
 44 <210> SEQ ID NO: 2  
 45 <211> LENGTH: 21  
 46 <212> TYPE: DNA  
 47 <213> ORGANISM: Artificial Sequence  
 49 <220> FEATURE:  
 50 <223> OTHER INFORMATION: Synthetic primer  
 52 <400> SEQUENCE: 2  
 53 cggaatcgcg ctacgcgctg g  
 55 <210> SEQ ID NO: 3  
 56 <211> LENGTH: 20  
 57 <212> TYPE: DNA  
 58 <213> ORGANISM: Artificial Sequence  
 60 <220> FEATURE:  
 61 <223> OTHER INFORMATION: Synthetic primer  
 63 <400> SEQUENCE: 3

21

## RAW SEQUENCE LISTING

DATE: 04/08/2002

PATENT APPLICATION: US/09/913,414

TIME: 14:15:05

Input Set : A:\08106-005001.TXT

Output Set: N:\CRF3\04082002\I913414.raw

```

64 gggagctcga agggtcctga
66 <210> SEQ ID NO: 4
67 <211> LENGTH: 555
68 <212> TYPE: PRT
69 <213> ORGANISM: Caulobacter crescentus
71 <400> SEQUENCE: 4
72 Met Phe Lys Arg Ser Gly Ala Lys Pro Thr Ile Phe Asp Gln Ala Val
73 1 5 10 15
74 Leu Val Ala Arg Pro Ala Val Ile Thr Ala Met Val Phe Ser Phe Phe
75 20 25 30
76 Ile Asn Ile Leu Ala Leu Val Ser Pro Leu Tyr Met Leu Gln Val Tyr
77 35 40 45
78 Asp Arg Val Leu Thr Ser Arg Asn Val Ser Thr Leu Ile Val Leu Thr
79 50 55 60
80 Val Ile Cys Val Phe Leu Phe Leu Val Tyr Gly Leu Leu Glu Ala Leu
81 65 70 75 80
82 Arg Thr Gln Val Leu Val Arg Gly Gly Leu Lys Phe Asp Gly Val Ala
83 85 90 95
84 Arg Asp Pro Ile Phe Lys Ser Val Leu Asp Ser Thr Leu Ser Arg Lys
85 100 105 110
86 Gly Ile Gly Gly Gln Ala Phe Arg Asp Met Asp Gln Val Arg Glu Phe
87 115 120 125
88 Met Thr Gly Gly Leu Ile Ala Phe Cys Asp Ala Pro Trp Thr Pro Val
89 130 135 140
90 Phe Val Ile Val Ser Trp Met Leu His Pro Phe Phe Gly Ile Leu Ala
91 145 150 155 160
92 Ile Ile Ala Cys Ile Ile Ile Phe Gly Leu Ala Val Met Asn Asp Asn
93 165 170 175
94 Ala Thr Lys Asn Pro Ile Gln Met Ala Thr Met Ala Ser Ile Ala Ala
95 180 185 190
96 Gln Asn Asp Ala Gly Ser Thr Leu Arg Asn Ala Glu Val Met Lys Ala
97 195 200 205
98 Met Gly Met Trp Gly Gly Leu Gln Ala Arg Trp Arg Ala Arg Arg Asp
99 210 215 220
100 Glu Gln Val Ala Trp Gln Ala Ala Ala Ser Asp Ala Gly Gly Ala Val
101 225 230 235 240
102 Met Ser Gly Ile Lys Val Phe Arg Asn Ile Val Gln Thr Leu Ile Leu
103 245 250 255
104 Gly Gly Gly Ala Tyr Leu Ala Ile Asp Gly Lys Ile Ser Ala Gly Ala
105 260 265 270
106 Met Ile Ala Gly Ser Ile Leu Val Gly Arg Ala Leu Ala Pro Ile Glu
107 275 280 285
108 Gly Ala Val Gly Gln Trp Lys Asn Tyr Ile Gly Ala Arg Gly Ala Trp
109 290 295 300
110 Asp Arg Leu Gln Thr Met Leu Arg Glu Glu Lys Ser Ala Asp Asp His
111 305 310 315 320
112 Met Pro Leu Pro Glu Pro Arg Gly Val Leu Ser Ala Glu Ala Ala Ser
113 325 330 335
114 Ile Leu Pro Pro Gly Ala Gln Gln Pro Thr Met Arg Gln Ala Ser Phe

```

20

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,414

DATE: 04/08/2002

TIME: 14:15:05

Input Set : A:\08106-005001.TXT

Output Set: N:\CRF3\04082002\I913414.raw

```

115          340          345          350
116 Arg Ile Asp Ala Gly Ala Ala Val Ala Leu Val Gly Pro Ser Ala Ala
117          355          360          365
118 Gly Lys Ser Ser Leu Leu Arg Gly Ile Val Gly Val Trp Pro Cys Ala
119          370          375          380
120 Ala Gly Val Ile Arg Leu Asp Gly Tyr Asp Ile Lys Gln Trp Asp Pro
121 385          390          395          400
122 Glu Lys Leu Gly Arg His Val Gly Tyr Leu Pro Gln Asp Ile Glu Leu
123          405          410          415
124 Phe Ser Gly Thr Val Ala Gln Asn Ile Ala Arg Phe Thr Glu Phe Glu
125          420          425          430
126 Ser Gln Glu Val Ile Glu Ala Ala Thr Leu Ala Gly Val His Glu Met
127          435          440          445
128 Ile Gln Ser Leu Pro Met Gly Tyr Asp Thr Ala Ile Gly Glu Gly Gly
129          450          455          460
130 Ala Ser Leu Ser Gly Gly Gln Arg Gln Arg Leu Ala Leu Ala Arg Ala
131 465          470          475          480
132 Val Phe Arg Met Pro Ala Leu Leu Val Leu Asp Glu Pro Asn Ala Ser
133          485          490          495
134 Leu Asp Gln Val Gly Glu Val Ala Leu Met Glu Ala Met Lys Arg Leu
135          500          505          510
136 Lys Ala Ala Lys Arg Thr Val Ile Phe Ala Thr His Lys Val Asn Leu
137          515          520          525
138 Leu Ala Gln Ala Asp Tyr Ile Met Val Ile Asn Gln Gly Val Ile Ser
139          530          535          540
140 Asp Phe Gly Glu Arg Asp Arg Cys Trp Pro Ser
141 545          550          555
143 <210> SEQ ID NO: 5
144 <211> LENGTH: 435
145 <212> TYPE: PRT
146 <213> ORGANISM: Caulobacter crescentus
148 <400> SEQUENCE: 5
149 Met Lys Pro Pro Lys Ile Gln Arg Pro Thr Asp Asn Phe Gln Ala Val
150 1          5          10          15
151 Ala Arg Ile Gly Tyr Gly Ile Ile Ala Leu Thr Phe Val Gly Leu Leu
152          20          25          30
153 Gly Trp Ala Ala Phe Ala Pro Leu Asp Ser Ala Val Ile Ala Asn Gly
154          35          40          45
155 Val Val Ser Ala Glu Val Ser Gln Asp Val Gln His Leu Glu Gly Gly
156          50          55          60
157 Met Leu Ala Lys Ile Leu Val Arg Glu Gly Glu Lys Val Lys Ala Gly
158 65          70          75          80
159 Gln Val Leu Phe Glu Leu Asp Pro Thr Gln Ala Asn Ala Ala Ala Gly
160          85          90          95
161 Ile Thr Arg Asn Gln Tyr Val Ala Leu Lys Ala Met Glu Ala Arg Leu
162          100          105          110
163 Leu Ala Glu Arg Asp Gln Arg Pro Ser Ile Ser Phe Pro Ala Asp Leu
164          115          120          125
165 Thr Ser Gln Arg Ala Asp Pro Met Val Ala Arg Ala Ile Ala Asp Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,414

DATE: 04/08/2002

TIME: 14:15:05

Input Set : A:\08106-005001.TXT

Output Set: N:\CRF3\04082002\I913414.raw

```

166      130      135      140
167 Gln Ala Gln Phe Thr Glu Arg Arg Gln Thr Ile Gln Gly Gln Val Asp
168 145      150      155      160
169 Leu Met Asn Ala Gln Arg Leu Gln Tyr Gln Ser Glu Ile Glu Gly Ile
170      165      170      175
171 Asp Arg Gln Thr Gln Gly Leu Lys Asp Gln Leu Gly Phe Ile Glu Asp
172      180      185      190
173 Glu Leu Ile Asp Leu Arg Lys Leu Tyr Asp Lys Gly Leu Val Pro Arg
174      195      200      205
175 Pro Arg Leu Leu Ala Leu Glu Ala Arg Ala Gly Ser Leu Ser Gly Ser
176      210      215      220
177 Ile Gly Arg Leu Thr Ala Asp Arg Ser Lys Ala Val Gln Gly Ala Ser
178 225      230      235      240
179 Asp Thr Gln Leu Lys Val Arg Gln Ile Lys Gln Glu Phe Phe Glu Gln
180      245      250      255
181 Val Ser Gln Ser Ile Thr Glu Thr Arg Val Arg Leu Ala Glu Val Thr
182      260      265      270
183 Glu Lys Glu Val Val Ala Ser Asp Ala Gln Lys Arg Ile Lys Ile Val
184      275      280      285
185 Ser Pro Val Asn Gly Thr Ala Gln Asn Leu Arg Phe Phe Thr Glu Gly
186      290      295      300
187 Ala Val Val Arg Ala Ala Glu Pro Leu Val Asp Ile Ala Pro Glu Asp
188 305      310      315      320
189 Glu Ala Phe Val Ile Gln Ala His Phe Gln Pro Thr Asp Val Asp Asn
190      325      330      335
191 Val His Met Gly Met Val Thr Glu Val Arg Leu Pro Ala Phe His Ser
192      340      345      350
193 Ala Gly Asn Pro Asp Pro Glu Arg His Asp Pro Val Ala Val Ala Asp
194      355      360      365
195 Arg Ile Ser Asp Pro Gln Lys Gln Ala Arg Leu Phe Leu Gly Ile Val
196      370      375      380
197 Arg Val Asp Val Lys Gln Leu Pro Pro His Leu Arg Gly Arg Val Thr
198 385      390      395      400
199 Ala Gly Met Pro Ala Gln Val Ile Val Pro Thr Gly Glu Arg Thr Val
200      405      410      415
201 Leu Gln Tyr Leu Phe Ser Pro Leu Arg Asp Thr Leu Arg Thr Thr Met
202      420      425      430
203 Arg Glu Glu
204      435
206 <210> SEQ ID NO: 6
207 <211> LENGTH: 100
208 <212> TYPE: PRT
209 <213> ORGANISM: Caulobacter sp. FWC1
211 <400> SEQUENCE: 6
212 Thr Thr Asp Thr Leu Lys Phe Ala Asn Thr Gly Thr Glu Thr Phe Thr
213 1      5      10      15
214 Ser Thr Lys Val Asp Leu Thr Gly Val Asn Asp Phe Thr Ala Ala Leu
215      20      25      30
216 Asn Ala Ala Ala Ala Gly Asn Gly Gly Gly Asn Gly Ile Ile Thr Trp

```

## RAW SEQUENCE LISTING

DATE: 04/08/2002

PATENT APPLICATION: US/09/913,414

TIME: 14:15:05

Input Set : A:\08106-005001.TXT

Output Set: N:\CRF3\04082002\I913414.raw

```

217          35          40          45
218 Phe Gln Tyr Gly Gly Asn Thr Tyr Ile Val Glu Asp Arg Asp Ala Gly
219      50          55          60
220 Asn Thr Phe Asn Val Ala Thr Asp Ile Val Val Lys Leu Thr Gly Ala
221 65          70          75          80
222 Val Asp Leu Ser Thr Ala Val Leu Ser Ala Phe Gly Arg Arg Ser Ser
223          85          90          95
224 Leu Thr Leu Val
225          100
227 <210> SEQ ID NO: 7
228 <211> LENGTH: 100
229 <212> TYPE: PRT
230 <213> ORGANISM: Caulobacter sp. FWC19
232 <400> SEQUENCE: 7
233 Arg Ala His Met Ile Leu Lys Pro Thr Arg His Val Ser Asp Arg Trp
234 1          5          10          15
235 Gly Arg His Val Ala Arg Leu Val Gln Leu Pro Gly Arg Pro Cys Pro
236      20          25          30
237 Lys Leu Ser Asp Ala Ala Thr Thr Gly Asn Ala Ser His Lys Val Ser
238      35          40          45
239 Trp Phe Val Tyr Gly Gly Asp Thr Tyr Leu Val Lys Met Ser Thr Leu
240      50          55          60
241 Ala Pro Pro Ser Lys Thr Ala Arg Thr Ile Val Val Lys Leu Thr Gly
242 65          70          75          80
243 Thr Thr Asn Asp Leu Thr Lys Ala Thr Phe Asp Gly Ala Ala His Thr
244          85          90          95
245 Leu Thr Leu Gly
246          100

```

**VERIFICATION SUMMARY**

**PATENT APPLICATION: US/09/913,414**

**DATE: 04/08/2002**

**TIME: 14:15:06**

**Input Set : A:\08106-005001.TXT**

**Output Set: N:\CRF3\04082002\I913414.raw**